Chapter 5: Multivariate Analysis of FSP Certification Costs and Error Rates

The analysis in this chapter seeks to expand the understanding of trends in error rates between 1989 and 2001 by adding a measure of certification cost to the explanatory variables used by previous studies. (In this discussion, "cost" refers to the expenditure of FSP administrative funds.) We first present the analysis objectives and the framework of assumptions about the relationship of costs to error rates. Next, we define the conceptual model for the analysis. In subsequent sections, we describe the data and methods used to model this relationship. The chapter concludes with a discussion of the results. The conclusions and limitations of the study are discussed in Chapter 6. Additional technical information on the methods is presented in Appendix C.

Analysis Objectives and Framework

The primary objective of this analysis was to estimate the impact of FSP certification costs on error rates, while controlling for other variables that affect the likelihood of error (and, therefore, the amount of effort needed to achieve a given level of accuracy).[1] The collection of data on costs, caseload characteristics, and error rates provided the opportunity to undertake this analysis. The framework for the analysis included hypotheses about the relationships of error rates to certification tasks, FSP workforce effort, Federal and State policies, and caseload characteristics.

Relationship of Certification Tasks to Error Rates

The outputs of FSP agencies differ in the accuracy of eligibility decisions, i.e., the positive and negative error rates as measured by the quality control (QC) process. Food stamp eligibility workers make **positive** errors when they approve benefits for ineligible households, or when they approve benefits that are greater than what households are eligible to receive. Workers make **negative** errors when they deny benefits to eligible households or approve benefits that are less than what households are eligible to receive. The appropriate measure of the overall level of accuracy for a FSP agency is a combination of positive and negative error rates.

When the number of households participating in the FSP (the size of the caseload) increases, the State FSP agency's workforce must accomplish more certification tasks. The number of tasks also increases when the caseload becomes more volatile, i.e., there are more entries of participants or more changes in participant circumstances affecting eligibility, relative to the size of the caseload. Each task that workers perform entails the potential for error, as does each change in circumstances to which workers do not respond (tasks that should be performed but are not, due to lack of information or lack of time). Some tasks have more potential for error than others, depending on the amount of information needed, the availability and reliability of the information, and the complexity of the decision rules.

FSP Workforce Time and Error Rates

The time that FSP agency eligibility workers spend on certification tasks is one important factor in determining the level of certification accuracy. For a given amount of time spent, workers with more

[1] We use the term "costs" to refer to expenditures that are allocated to the Food Stamp Program.

certification tasks to accomplish may spend less time on each task and, consequently, make more errors. The likelihood of error for a given amount of worker time is also likely to be greater if workers have a more difficult set of tasks to accomplish. (Difficulty may be a function of the risk of error or the steps required to complete the task.) Conversely, when the volume of tasks or their difficulty is less relative to amount of time spent, workers can devote more time to making sure that each case decision is accurate.

This relationship of certification worker time to accuracy only holds if other factors are held constant, however. When workers have less time available relative to the volume and difficulty of the tasks, they may maintain the same level of accuracy by processing applications on a less timely basis. Additional certification worker hours may be used for purposes other than error reduction, such as to improve timeliness or access (e.g., keeping offices open longer hours or out-stationing workers at locations other than FSP offices).

The amount of front-line eligibility worker hours is not the only human resource that may affect certification accuracy. Additional resources for supervision, training, and promoting worker morale and teamwork may be expected to reduce the level of errors. Quality control, management evaluation, and fair hearings can affect error rates through feedback about process improvements that are needed.

For the purposes of this analysis, we adopted an inclusive definition of "certification-related" costs that included both the certification costs identified in FSP reports and other costs that were expected to have an impact on certification errors. Our definition of certification-related costs included the group of cost reporting categories labeled as "miscellaneous", which consisted primarily of costs for quality control, management evaluation, verification of alien eligibility, and fair hearings.[2] In addition, the "unspecified other" category was included, because at least part of these costs was likely to be related to certification and error-reduction (as discussed in Chapter Two). FSP functions were excluded from the definition of certification-related costs if they were not expected to affect QC errors. Benefit issuance is a separate process with its own measures of accuracy; EBT was implemented in part to reduce errors and fraud in this process. Food Stamp Employment and Training (FSE&T) and Food Stamp Nutrition Education (FSNE) services address program goals other than minimizing error.

Conceptual Model of FSP Administrative Costs and Errors

We began the analysis with the following conceptual model for the relationship of FSP administrative costs to errors:

$$\text{ERROR}=f(\text{CERTCOST},X,\ \text{POLICY},\text{FP}) \qquad (1)$$

where ERROR is a weighted index combining positive and negative error rates in a State in a fiscal year, CERTCOST is the corresponding certification-related expenditure per FSP household, X is a vector of caseload characteristics, POLICY is a vector of policies determining the actions taken to prevent and detect errors, and FP is a vector of factor prices determining the effective output per dollar of CERTCOST. The factor prices include wage rates, employee benefit prices, and prices for goods and services needed to support certification-related labor.

---

[2] Outreach and research and demonstration projects were also included in this category, but they were relatively minor components.

The model represented by equation (1) treats the amount of funding that a State allocates to certification activities as exogenous. This seems like a justifiable assumption because State FSP administrative budgets are fixed prior to the start of the fiscal year. Of course, over time, FSP administrative budgets are endogenous because States increase these budgets as workload increases and decrease them as workload falls. Furthermore, agencies have some latitude to augment FSP administrative budgets by reallocating funds from other programs. Nevertheless, treating FSP administrative budgets and the portion allocated to certification activities as fixed in the short run is a useful assumption for the econometric modeling that is not seriously discrepant from the reality of FSP administration.

A possible problem with the model in equation (1) is that the cost per FSP household in one year may be influenced by the error rates in a preceding year, particularly when the State incurs financial sanctions for excessive error rates. We address this possibility by using modeling approaches that allow for a lagged effect of past error rates, as described later in the chapter.[3]

The general conceptual model presented above relates certification-related costs and factor prices to error rates, but for this analysis we used an alternative formulation:

$$\text{ERROR}=\text{F}(\text{EFFORT},\text{X},\text{POLICY}) \qquad (2)$$

where EFFORT is the quantity of administrative resources expended on certification-related activities. (All variables in this equation are State-level annual measures, so State and year subscripts are implied.) Two agencies with the same certification-related cost would have different levels of administrative effort, if one had higher factor prices. The agency with the higher level of administrative effort would be expected to have a lower overall error rate, assuming that other factors affecting error rates were the same. The advantage of this formulation is that it creates a single variable that combines the effects of the certification-related cost and factor price (FP) variables.

Ideally, administrative effort would be measured as a vector of factor costs divided by the prices of those factors, yielding estimates of labor hours, units of computer processing services, square feet of space etc. Such analysis would require much more detailed cost data than are available from FNS, which only receives data on the total cost of each program function.

Therefore, the analysis used a scalar measure of administrative effort:

$$\text{EFFORT}=\text{CERTCOST}/\ W_{\text{FTE}} \qquad (3)$$

where CERTCOST is the annual certification-related cost per FSP household (as in equation 1) and W_{FTE} is the annual public welfare worker wage rate per full-time equivalent employee. (Again, year and State subscripts are implied.) The resulting effort measure was a proxy for the quantity of administrative resources per food stamp household. The cost measure included allocated overhead costs (such as facilities, supplies, non-ADP equipment, and ancillary services). Thus, the effort measure cannot be interpreted as labor alone, but rather labor with a multiplier for overhead costs.[4]

3 Past error rates may also influence state-level policies and procedures such as change reporting requirements and certification period lengths.

4 For example, assume that overhead costs are allocated to the FSP by adding a fixed amount (O) per full-time equivalent worker. Thus, $\text{CERTCOST}=\text{FTE}*(W_{\text{FTE}}+\text{O})$, and $\text{EFFORT}=\text{FTE}*(1+\text{O}/\ W_{\text{FTE}})$. Actual cost allocation procedures may be more complex.

Allocation of Certification Costs to the FSP

A complication arises because of the way that certification costs are determined. The total certification cost for a State is the sum of the cost of certifying FSP-only households and the FSP's allocated share of certification costs for FSP households receiving other State-administered benefits, such as AFDC or TANF and General Assistance. The allocated FSP share is determined by each State's cost allocation plan, which may vary from other States but must receive Federal approval. Under cost allocation rules that applied before PRWORA, some shared costs for FSP/AFDC cases were allocated to the AFDC program as the "primary program". Under Public Law 105-185, enacted in 1997, States were required to prorate these costs between TANF and the FSP. Thus, a State's FSP certification effort and cost could increase even if its total certification cost for all programs did not change.[5] The models used in this study controlled for this discontinuity, as discussed later in this chapter.

Automated Data Processing Costs

In addition to human resources, States use automated data processing systems to prevent and detect certification errors. As noted in Chapter Two, the level of cost per FSP household for developing and operating these systems fell from 1989 to 1994 and increased thereafter (although the ADP development cost per FSP household peaked in 1999 and the ADP operations cost per FSP household peaked in FY2000). Thus, a complete model of the impact of FSP resources on error rates should take ADP spending into account.

The computation of the effort measure excluded automated data processing (ADP) costs from CERTCOST, because it was clearly inappropriate to treat the ratio of ADP costs to wages as an estimate of data processing units. Reasoning that factor prices for data processing are largely set in a national (or even international) market, we preferred to treat the ADP cost per FSP household as a separate independent variable.

The models ultimately did not include this variable, however, due to difficulties encountered in the analysis. In the exploratory phase of this analysis, we used several different measures of ADP development and operating costs per FSP household as alternative independent variables. The resulting positive coefficient –implying that increased ADP spending contribute to an increase in the error measure—was contrary to expectations. One possible interpretation was that the ADP cost measures were proxying for an omitted variable; another was that errors increased in the short run with new ADP systems, because of implementation challenges. We also recognized the fundamental problem that ADP development spending represented an investment, and that a State's level of automation could be more closely related to cumulative ADP spending over the history of the FSP than on the spending for a specific period. For example, agencies that invested less in ADP systems during the 1980's might have spent more to catch up with their peers in the 1990's. A further factor that may have affected the observed relationship of ADP spending to error rates was the fact that many States had to test and modify

[5] TANF replaced the open-ended matching of administrative costs and benefits with a block grant. Thus, there was, in theory, an incentive for States to shift costs away from TANF and toward the FSP, which remained an entitlement. Concerns about this possibility were a factor behind the cost allocation provision in P.L. 105-185 (Carmody and Dean, 1998). No evidence of such a shift was gathered for this study, but if it had occurred, it would have further affected the comparability of certification costs between the pre-PRWORA and post-PRWORA periods.

or renovate their ADP systems in preparation for the year 2000. Lacking confidence in the validity of the ADP cost measures for this analysis, we chose to leave them out and rely on other methods discussed below to assure that this omission did not bias the results.

Certification Cost and FSP Agency Performance

While certification accuracy has historically been the primary measure of the performance of State FSP agencies, there are other dimensions of performance that may be affected by the level of certification effort. These dimensions include timeliness and accessibility. As discussed below, these are important dimensions, and States may direct incremental certification effort to improving performance on these dimensions rather than to improving certification accuracy.

The timeliness of application processing has always been an FSP performance indicator. Unlike error rates, the standard of performance is absolute: FSP agencies are required to act on applications within 30 days (with some exceptions). This performance measure was not feasible to analyze, because State-level time series data on the timeliness of all applications, including denied applications, are not published.

In recent years, increased attention has focused on the level of accessibility and customer service. The shortening of certification periods in the late 1990’s gave rise to concern that increased burdens on FSP participants were discouraging participation or at least changing participants’ views of the relative costs and benefits of participation. For example, Kornfeld (2002) found that the increased proportion of households with short certification periods contributed to the decline in FSP participation among households with earnings. Bartlett and others (2004) identified a number of dimensions of accessibility that were related to the probability that a household would complete an application. Under the 2002 Farm Security and Rural Investment Act, States with high or improved rates of FSP participation can receive bonuses.

If the impact of certification effort on FSP agency performance is multi-dimensional, a model of the impact of certification effort on all dimensions of performance might be desirable, but this approach was not feasible. As noted above, data on the timeliness of applications were not available. We were unable to identify any State-level longitudinal data on FSP policies affecting accessibility for the study period. Furthermore, a model of the effects of certification effort on both error and participation rates would entail simultaneous equations or instrumental variables. Given these issues, and the focus of FSP administration on error rates as the primary measures of performance during the study period, we did not attempt to take accessibility into account. We acknowledge that the analysis represents a simplification of the outputs of the FSP, and that future analyses may need to revisit this problem.

Data

This analysis used a panel of the 50 States and the District of Columbia over the 13 years from 1989 to 2001, in order to examine the relationship between administrative effort per food stamp household and food stamp error rates.

Dependent Variable

The dependent variable used in this analysis was an index of error computed as a weighted sum of annual positive error rates and negative error rates. This approach was consistent with FSP policy, which recognizes the importance of minimizing both types of errors and combines them in the payment error rate on which sanctions are based. Positive error rates were calculated by aggregating two types of overpayment errors—the percentage of total FSP cases receiving benefits greater than statutorily prescribed levels by at least $25 per month, and the percentage of active FSP cases that were not eligible to receive any benefits under program rules. Negative error rates were calculated by aggregating two types of underpayment errors—the percentage of total FSP cases receiving benefits less than statutorily prescribed levels by at least $25 per month, and the ratio of negative action errors to total caseload. As discussed in the preceding chapter, the denominator for all these rates was the sum of active cases and cases subject to negative action (denial, suspension, or closure).

The error index, ERROR, was calculated as:

$$ERROR = ERROR_p + l\,ERROR_n \tag{4}$$

where $ERROR_p$ is the positive error rate, $ERROR_n$ is the negative error rate, and λ is a parameter representing the relative difficulty of eliminating negative versus positive errors, estimated via grid search as described in Appendix C. The mean State positive error rate during the period 1989-2001 was 9.3 percent and the mean negative error rate was 4.4 percent.[6] The error index depends on the value of λ. We introduced the parameter λ into the model because while spending more on the administration of food stamps was expected to reduce errors, positive errors may be more or less difficult to affect than negative errors. We estimated that $\lambda = 1.45$ with a standard error of 0.16, implying that the amount of resources required to reduce the positive error rate by 1 percentage point would reduce the negative error rate by 0.69 (=1/1.45) percentage points (assuming that only one rate changes at a time). For example, assume a simple model such that ERROR=A-.001(RESOURCES). Thus, an additional 10 units of resources would reduce ERROR by .01. If $ERROR_n$ stays constant, $ERROR_p$ is reduced by 1 percentage point; if $ERROR_p$ stays constant, $ERROR_n$ is reduced by 0.69 percentage points. Conversely, the amount of resources required to reduce negative error by one percentage point are 1.45 times the resources required to reduce positive error by the same amount (again holding one error rate constant while the other changed).[7]

Figure 33 shows the national trend in the error index in the context of the national trend in FSP households. As the figure shows, the error index increased from 1989 to 1992, and then decreased in 1993 through 1996, increased in 1997 and 1998, then fell in 1999 to 2001. Although the overall trends suggested a positive association between the error index and the number of FSP households, the fluctuations in the error index trends suggest that they were influenced by other factors as well.

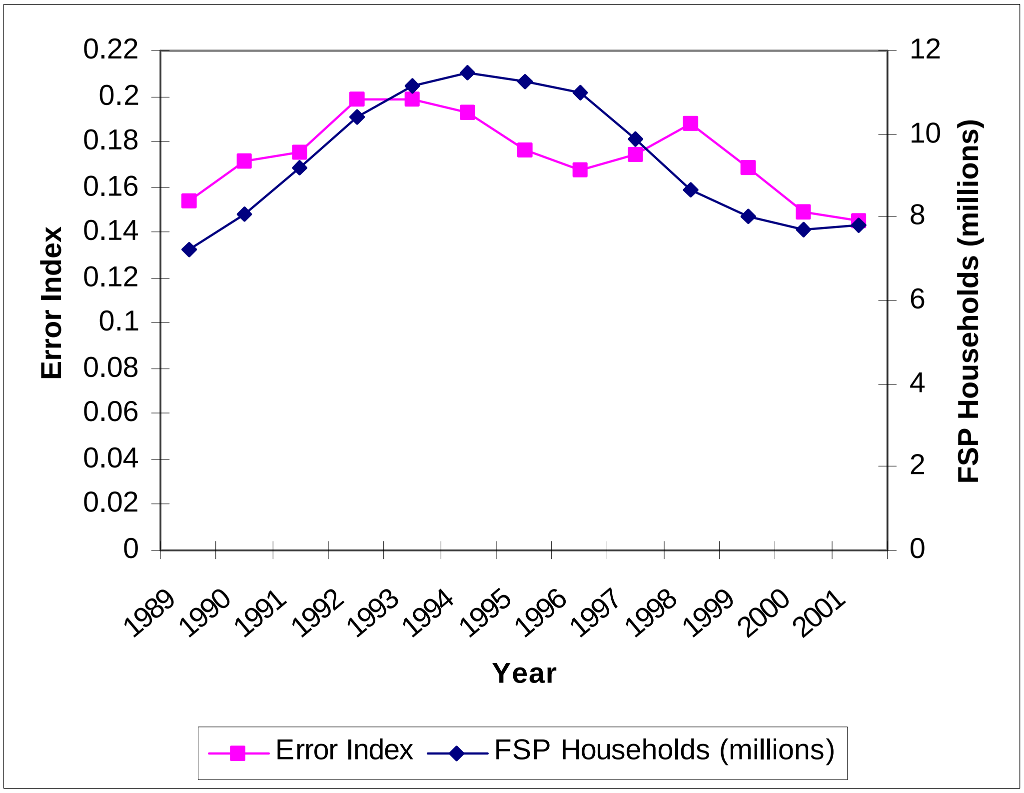
Effort

The main focus of this analysis was the impact of a change in administrative effort on the error index. The ideal measure of effort was the ratio of full-time food stamp workers to food stamp households.

6 We remind the reader that these do not correspond to published rates, because of our use of a common denominator for all types of errors.

7 The estimate of λ was robust to choice of model specification, i.e., its value did not change materially when variables were added to or subtracted from the model. The parameter and its standard error were estimated for the fixed effects model described first.

Figure 33
National Trends in Error Index and FSP Households, 1989-2001



Note: See text for definition of error index.

Because we could not observe this variable directly, we used as a proxy effort measure the certification-related cost per FSP household, normalized by dividing the cost by the state wage for a full-time public welfare worker. Thus, the effort measure was computed as in (5) below:

$$EFFORT = \left[\frac{ATC - AADPC - AIC - AFSNEC - AETC}{HH}\right]\left[\frac{1}{W_{FTE}}\right] \qquad (5)$$

where ATC is the annual total FSP cost, AADPC is the annual automated data processing (ADP) cost, AIC is the annual issuance cost, AFSNEC is the annual Food Stamp Nutrition Education cost, AETC is the annual employment and training cost, HH is the number of food stamp households (computed by averaging monthly data) and W_{FTE} is the annual public welfare worker wage rate per full-time equivalent (FTE) employee. All variables are specific to a year within a State, and costs are in 2001 dollars. The first quantity in equation (5) represents CERTCOST (as discussed in the preceding section), with the numerator representing the total costs of certification and other related activities to manage and assure FSP eligibility. Dividing CERTCOST by the public welfare wage rate normalized the effort measure to control for differences in pay rates.

Other Independent Variables

Although the main focus of this analysis was on the impact of effort on the error index, we included other observable covariates in the analysis for two purposes. First, the error index may vary for different case types or under different program conditions (rules etc.), holding resources constant. Second, the level of effort required to process cases with a given level of accuracy may vary by case type or program conditions. Therefore, we introduced control variables intended to control for time-varying differences in caseload characteristics and program conditions.

PRWORA is an indicator for the years when the Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA) was in effect. We considered two ways that program conditions under PRWORA could affect the level of FSP errors while effort and other factors were held constant (i.e., affect the intercept of the regression line). On the one hand, the implementation of PRWORA had the potential to destabilize operations, because workers might be more focused on learning new rules and on getting clients employed. Furthermore, PRWORA changed the relationship between food stamps and cash assistance to families, so that certification and case management for FSP/public assistance households became more complicated and thus more error-prone. On the other hand, FSP agencies made changes in staff training, management, mission, and incentives for performance with the implementation of PRWORA, and these changes had the potential to improve the accuracy of certifications and other caseworker actions. In our sample, 38.5 percent of observations were from the post-PRWORA period (1997-2001).

PEFFORT was defined as the product of *PRWORA* and *EFFORT.* We included this variable to account for the possibility that changes in the FSP during the post-PRWORA period might alter the effect of effort on error (i.e., a change in the slope of the regression line). As described elsewhere in this report, cost allocation rule changes effectively increased the FSP's share of the costs of serving households receiving both food stamps and cash assistance. As a result, we expected to find that an additional unit of effort was less effective at reducing the error index in the post-PRWORA period.[8] This variable also accounted for the possibility that other changes in FSP operations associated with PRWORA implementation might increase or decrease the impact of a given amount of effort on the error index (e.g., changes in efficiency as a result of adaptation to new rules or re-engineering).

TANF is the percent of food stamp households receiving Aid to Families with Dependent Children (AFDC) or its successor, Temporary Aid for Needy Families (TANF). These households were expected to be less error-prone than other food stamp households, because the food stamp agency is required to have authoritative information on AFDC/TANF benefits, and these households are less likely to have earnings or other sources of income. In addition, the level of error with a given level of effort allocated to the FSP was expected to be less in a State with a high percentage of FSP households receiving AFDC or TANF than in a State with a low percentage of FSP households receiving AFDC or TANF, because of the sharing of costs between the FSP and the AFDC/TANF program, as previously discussed.

PTANF was defined as the interaction of the TANF variable with the PRWORA variable. This variable was included to account for a possible differential impact of PRWORA on FSP operations, depending on the size of the TANF/food stamp caseload relative to the total food stamp caseload. In the pre-PRWORA period, 37.7 percent of food stamp households received AFDC benefits. In the post-PRWORA period,

[8] The PRWORA indicator variable was not expected to capture any of the effect of cost allocation changes, because the effect of those changes would vary according to the level of effort per FSP household.

24.9 percent of food stamp households received TANF benefits. PTANF also had the potential to pick up an indirect effect of changes in cost allocation rules on error: if more of the actual effort for FSP/TANF cases was allocated to the FSP under PRWORA, States with high percentages of FSP households on TANF would tend to have higher levels of error for a given level of reported (i.e., allocated) effort.

EARNINC is the percent of food stamp households with earned income in their case records.[9] Households with reported earnings are likely to have more volatile income and thus be more prone to underpayment or overpayment error. In our sample, the mean percentage of food stamp households with earned income was 24.6 percent.

SSINC is the percent of food stamp households with Social Security Old Age, Disability, and Survivors Insurance (OASDI) or Supplemental Security Income (SSI) benefits. These households were expected to be less error-prone than other food stamp households because the Food Stamp program can easily and definitively verify OASDI and SSI benefits through well-established data exchange systems. These households are also unlikely to have unreported earnings, a potential source of error that is not captured by the reported earnings indicator. The mean was 36.8 percent of food stamp households receiving OASDI or SSI benefits.

SINGLEPAR is the percent of food stamp households with children headed by a single adult. These households were expected to be less error-prone than households with two parents (and thus two potential earners), after controlling for the presence of any reported earnings, because they were less likely to have unreported earnings. The mean percentage of food stamp households with one or more children that had a single adult was 72.3 percent.

PCTEBT is the percent of food stamp households that receive benefits via electronic benefits transfer (EBT). Under the coupon issuance system, food stamp case workers dealt with replacement of lost or stolen coupons, but comparable functions under EBT are mainly handled by separate customer service centers. Therefore, greater use of electronic benefits transfer, relative to coupons, was expected to reduce interruptions that might contribute to case worker error. The first statewide implementation of EBT (in Maryland) was completed in 1993; by 2001 most States had implemented EBT. Over the 13 study years, the mean percentage of food stamp households receiving electronic benefits was 23.3 percent.[10]

FYUN0 is the state-specific unemployment rate. When unemployment rates are low, food stamp recipients are more likely to be employed and therefore subject to error due to fluctuations in employment and earnings. The mean State unemployment rate from 1989 to 2001 was 5.3 percent.[11]

---

9 This variable does not count households with only unreported earnings, so it understates the proportion of households with the potential for erroneous information on earnings.

10 In most years, PCTEBT was either 0 or 1. Values between 0 and 1 occurred during the transition from coupon to EBT issuance, and when certain States issued a portion of benefits in cash under special waivers.

11 A related variable, the change in the unemployment rate, was also used in alternate specifications. The results were similar to those presented here. An increase in the unemployment rate could increase error rates, because there would be more first-time, short-term food stamp recipients. These recipients might be more prone to error because of having no history of dealings with welfare workers. The change in unemployment rate was not significant when included with the unemployment rate in the model.

Finally, *CM13* is the percentage of food stamp cases with certification periods of one to three months. Short certification periods were expected to reduce error rates because more frequent reviews of eligibility. Analysis by the Center on Budget and Policy Priorities provided suggestive evidence in support of this hypothesis (CBPP, 2001). The mean percentage of food stamp households that have short certification periods is 9.9 percent.[12]

The means and standards deviations of the dependent variable and the independent variables, along with their definitions, are found in table 8.

Methods

We estimated four models to test the association between effort and the food stamp error index—a simple fixed effects model, a fixed effects model that corrects for first-order autocorrelation and heteroskedasticity, a simple partial adjustment model, and a dynamic model using an Arellano-Bond estimator. The models are described below. Derivations of equations and other details are provided in Appendix C.

Fixed Effects Model

Our data are a panel of 50 States plus the District of Columbia over 13 years. We use a fixed effects model to estimate equation (6):

$$ERROR_{it} = \alpha_i + t_i'\delta_i + EFFORT'_{it}\beta_1 + PEFFORT_{it}'\beta_2 + X_{it}'\gamma + \varepsilon_{it} \qquad (6)$$

where $ERROR_{it}$ is the error index in State i at time t, α_i is a time-invariant state-effect, t_i is a state-specific linear time trend, δ_i is the state-specific coefficient on that linear time trend, β_1 is the parameter estimate on EFFORT, β_2 is the parameter estimate on PEFFORT, X_{it} is a row vector of control variables, and γ is a column vector of parameters conformable with X.[13]

The time-invariant state effect, α_i, controls for unmeasured static factors that vary across States. For example, urban States may have higher error rates than rural States, in which case α_i would control for urbanicity.[14] If those state effects were excluded and the omitted variables were correlated with EFFORT, the estimate of β_1 would be biased and inconsistent. The time effect, δ_i, controls for state-invariant

12 It is possible that assigning more cases of one type to shorter certification periods could lead to more errors in other types of cases, particularly when EFFORT is held constant. We did not view this as a zero-sum situation, however, instead expecting that increased use of short certifications was likely to be indicative a broader set of policies intended to reduce errors.

13 Alternate fixed effects model specifications were considered, including those with non-linear EFFORT terms, a log-linear functional form, and a log-log functional form. We found no evidence that any of these non-linear models provided a better fit for the data than the linear model we present. For example, the inclusion of a squared EFFORT term did not improve the model; the coefficient on this variable was never statistically different from zero. Thus, we concluded that over the range of error rates that we are examining, there exists a linear relationship between EFFORT and ERROR.

14 In a cross-sectional analysis, Puma and Hoaglin (1987) found that the incidence and amount of overpayments were positively related to a State’s population density. This variable was not expected to vary greatly over time within a State, so it was not included as a separate variable in the models.

Table 8

Means and Standard Deviations of Analysis Variables

Variable	Definition	Mean (S.D.)
ERROR	Weighted total error rate	0.157 (0.048)
EFFORT	Certification-related cost per FSP household, normalized by the state wage for a full-time public welfare worker	0.010 (0.004)
PRWORA	Indicator for post-PRWORA period (1997-2001)	0.389 (0.488)
PEFFORT	Interaction between EFFORT and PRWORA	0.005 (0.006)
TANF	Percent of food stamp households receiving AFDC or TANF	0.325 (0.126)
PTANF	Interaction between TANF and PRWORA	0.097 (0.141)
EARNINC	Percent of food stamp households with earned income in case record	0.247 (0.080)
SSINC	Percent of food stamp households with OASDI or SSI benefits	0.370 (0.104)
SINGLEPAR	Percent of food stamp households with children headed by a single adult	0.722 (0.087)
PCTEBT	Percent of food stamp households that receive electronic benefits	0.236 (0.395)
FYUN0	Unemployment rate	0.053 (0.015)
CM13	Percent of food stamp households with 1-3 month certification periods	0.100 (0.129)
n		654[a]

[a] Negative action error data were unavailable for 9 observations.

unmeasured factors that have linear trends over time.[15] For example, unmeasured improvements in data processing technology may affect effort rates over time, and linear time trends take this effect into account. With these two types of fixed effects in the model, the remaining control variables account for variation in error rates attributable to within-state changes in measured factors. Factors that do not vary over time are already taken into account by α_i. Additionally, factors that vary linearly over time are already taken into account by δ_i. The parameters associated with the control variables capture whatever remaining partial correlation exists between error rates and the measured variables.[16]

15 We also estimated models that included national time effects. These models produced results that were consistent with the findings we present here using state-specific linear time trends. Thus our results are robust to the specification of the time effect.

16 We considered an alternate specification with random state effects, which would have the advantage of being more efficient if the null hypothesis of no systematic difference between fixed and random effects coefficients were true. We ran a Hausman test, however, and reject the null hypothesis at $p < 0.001$. Thus, a fixed effects specification appears more appropriate.

We assumed that the disturbance term is independently and identically distributed with zero mean and constant variance. Thus, the model described by equation (3) assumes that there is no heteroskedasticity or autocorrelation of residuals. In the following model, we relaxed these assumptions.

Prais-Winsten FGLS Models

Autocorrelation and heteroskedasticity are frequent problems when analyzing panel data. Autocorrelation can arise for two reasons. First, if both past and present values of some explanatory variables affect the dependent variable—and these lagged variables are omitted from the model—then the resulting disturbance term may reflect a systematic pattern due to serial correlation across periods. Second, if the dependent variable in period *t* is not independent of the dependent variable in period *t-1*, then the process itself will have an autocorrelated error structure. Heteroskedasticity can arise because the error rate variable may have a higher variance in some States than in others; for example, the variance could be correlated with the size of the caseload.[17] If ignored, both autocorrelation and heteroskedasticity lead to biased and inconsistent standard errors and, consequently, misleading hypothesis tests and confidence intervals.

The Prais-Winsten estimate is a Feasible Generalized Least Squares (FGLS) estimator that works to “sweep” first-order autocorrelation from the model. The estimator requires two steps. First, using the fixed effects model described in equation (3), the analyst estimates the regression residuals. Since these are consistent measures of the error term (e_{it}), they are used to estimate the autocorrelation coefficient ρ. The estimate of ρ is used to transform the dependent variable and every independent variable, as described in the appendix.

The Prais-Winsten model was estimated in two ways—first, with a common estimate of the ρ parameter across States, and second with a state-specific estimate of ρ. Relative to a Prais-Winsten model using a common ρ, a model using state-specific estimates of ρ may produce less biased estimates if the autocorrelation parameters are not equal across States. It may be less efficient, however, because it requires additional parameter estimates. Because of this trade-off, we present and compare results using both Prais-Winsten models.

Partial Adjustment Model

The previous models treat the availability of resources as exogenous to the error index. A dynamic model, however, might better explain the data generating process. One form of dynamic model that we employed is the partial adjustment model.

The partial adjustment model assumes that States adjust their resources so as to achieve a desired level of errors, but only make these adjustments gradually, closing part of the gap between the actual and target error index each year. Details on the calculation of the parameters and their standard errors are found in Appendix C.

The simple partial adjustment model estimator may have several problems. First, if the residuals are autocorrelated, the estimated standard errors may be inconsistent, leading to misleading hypothesis

---

17 QC samples for all States are designed to achieve similar levels of precision, so the level of sampling error not likely to vary by size of State. Larger States, however, may differ from smaller States in other ways (for example, the heterogeneity of the food stamp caseload), which may produce more variability in error rates.

testing. Second, if the time series does not satisfy stationarity, coefficient estimates will be biased and inconsistent. When this model was estimated, however, the process was not found to be autocorrelated, and it was clearly stationary.

More importantly, the partial adjustment model is biased because the lagged value of the dependent variable is necessarily correlated with the error term. The model is consistent as the number of time periods approaches infinity, but with a period of just 13 years, we expect a bias on the order of 1/13 (Baltagi, 1995, p. 126). Although we might be willing to accept this bias, the Arellano-Bond approach provides an alternative.

Arellano-Bond Dynamic Model

Arellano and Bond (1991) used instrumental variables to surmount the problem of bias and inconsistency introduced when using the lagged dependent variable as a regressor. This method uses dependent variables lagged two and three periods as instruments for the one-period lagged dependent variable. The instruments are obtained in a dynamic panel model by using orthogonality conditions between lagged values of ERROR and the disturbances, v_{it} (Baltagi, 1995). The model is estimated via generalized method of moments (GMM) using the STATA command xtabond.[18]

Using the Arellano-Bond model has two benefits. First, the model imposes no distributional assumptions on the residuals, but only requires the absence of serial correlation. Second, it uses an instrumental variable method to account for the fact that one-period lagged values of the dependent variable will be correlated with the residuals. A drawback of the Arellano-Bond estimator, however, is that it requires three years of data to be dropped to utilize the instruments. The loss of information results in less efficient parameter estimates. Consequently, the Arellano-Bond estimator is not clearly preferable to the simple partial adjustment model. The latter may have a smaller mean-squared error.

Each of the four models presented above—the fixed effects model, the Prais-Winsten FGLS model, the partial adjustment model, and the Arellano-Bond model—makes different assumptions about the data generating process. As described above, each of these models has advantages and disadvantages. To the extent that each model provides a similar answer with respect to the impact of effort on food stamp error rates, one may conclude that the estimates are robust with respect to the choice of model.

Elasticities

In order to provide a unit-free measure of the impact of effort on food stamp errors, we calculated effort elasticities with respect to error. The elasticity is equal to the percentage change in ERROR resulting from a one percent increase in EFFORT, holding other variables constant. We calculated effort elasticities for both the pre-PRWORA period (1989-1996) and the post-PRWORA period (1997-2001), because both our expectations and the analysis indicated that the impact of EFFORT on ERROR was different in these two periods. This difference was represented by the PEFFORT variable, which was factored into the post-PRWORA elasticity along with the EFFORT variable. The calculation of elasticities in the pre-and post-PRWORA periods, along with their respective standard errors, which are not straightforward, can be found in Appendix C.

[18] Details of the model are provided in Baltagi (1995) and in Appendix C.

Results

The results of the five model specifications described above are presented in table 9. Each of the columns denotes the model used to generate the estimate, and each row represents a right-hand side variable of interest. In each case, the dependent variable is the error index. Both versions of the Prais-Winsten model are presented: the model estimating state-specific values of ρ (2a) and the model estimating a constant autocorrelation coefficient for all States (ρ) (2b). For the partial adjustment and Arellano-Bond models, the long-run parameter estimates are presented so as to allow direct comparison with the fixed effects and Prais-Winsten models.

Effort

The estimation results present a convincing case that, as expected, there is a strong negative association between the effort level put forth by States and the food stamp error index. The coefficient for EFFORT is estimated with a high degree of confidence (p <0.01) in all models. Thus, the results support the expectation that increased effort (as proxied by EFFORT) reduces error. The estimated coefficient for EFFORT is fairly similar across the models, with most values between –4.64 and –5.73, providing further evidence of the robustness of these results. The Arellano-Bond model does, however show a noticeably higher coefficient than the other models.

Nearly all of the models indicate a positive and significant association between PEFFORT (the interaction of the post-PRWORA period with EFFORT) and ERROR. The magnitude of the long-run effect of PEFFORT is quite similar across the models (2.80 to 3.44). With the exception of the Arellano-Bond model, the coefficient for PEFFORT is significantly greater than zero, with a high level of significance (p < 0.01) for the fixed effects model and both versions of the Prais-Winsten FGLS model. (The Arellano-Bond model generally has larger standard errors due to the limitations discussed in the preceding section.) Taken together with the findings for EFFORT, these results imply *a smaller impact of worker effort on error in the post-PRWORA environment*.

While it is important to confirm that higher levels of effort are associated with a lower error index, it is also important to consider the magnitude of the relationship. We used the model parameters to estimate the size of the relationship, expressed in the form of an elasticity. Table 10 presents the estimates of the elasticity of the error index with respect to effort for the pre- and post-PRWORA periods. The estimated elasticity for the pre-PRWORA period ranged from-0.276 to –0.377. Thus, holding the negative error rate constant, a 10 percent increase in effort reduced the positive error rate by 2.76 to 3.77 percent. Alternatively, with the positive error rate held constant, a 10 percent increase in effort reduced the negative error rate by 1.90 to 2.60 percent (reflecting the weighting in the error index such that an increment of effort that produced 1 percentage point of change in positive error rates would produce 0.69 percentage points' change in the negative error rate).

.

Table 9

Models of Association Between Effort and Error Index[a]

	Definition	(1) Fixed Effects	(2a) Prais-Winsten FGLS (ρ_i)[b]	(2b) Prais-Winsten FGLS (ρ)[b]	(3) Partial Adjustment[a]	(4) Arellano-Bond[a]
EFFORT	Certification-related cost per FSP household, normalized by the state wage for a full-time public welfare worker	-5.14*** (0.880)	-4.64*** (1.01)	-5.06*** (0.991)	-5.73*** (1.21)	-7.18*** (2.21)
PRWORA	Indicator for post-PRWORA period (1997-2001)	-0.062*** (0.016)	-0.049*** (0.015)	-0.060*** (0.018)	-0.052** (0.020)	-0.041 (0.035)
PEFFORT	Interaction between EFFORT and PRWORA	3.44*** (0.986)	2.92*** (0.844)	3.23*** (0.941)	2.80** (1.30)	2.78 (2.34)
TANF	Percent of food stamp households receiving AFDC or TANF	-0.023 (0.035)	-0.007 (0.031)	-0.021 (0.063)	-0.044 (0.046)	-0.064 (0.083)
PTANF	Interaction between TANF and PRWORA	0.136*** (0.034)	0.109*** (0.037)	0.128*** (0.041)	0.155*** (0.043)	0.136* (0.076)
EARNINC	Percent of food stamp households with earned income in case record	0.149*** (0.053)	0.200*** (0.052)	0.161*** (0.041)	0.235*** (0.067)	0.440*** (0.120)
SSINC	Percent of food stamp households with OASDI or SSI benefits	-0.112** (0.047)	-0.095** (0.043)	-0.101** (0.047)	-0.140** (0.063)	-0.151 (0.111)
SINGLEPAR	Percent of food stamp households with children headed by a single adult	-0.013 (0.004)	-0.004 (0.036)	0.004 (0.040)	0.012 (0.055)	-0.135 (0.094)
PCTEBT	Percent of food stamp households that receive electronic benefits	0.003 (0.006)	0.001 (0.005)	0.002 (0.005)	0.009 (0.007)	0.007 (0.013)
FYUN0	Unemployment rate	-0.185 (0.160)	-0.097 (0.194)	-0.135 (0.214)	-0.123 (0.218)	0.438 (0.376)
CM13	Percent of food stamp households with 1-3 month certification periods	-0.127*** (0.018)	-0.121*** (0.023)	-0.120*** (0.016)	-0.144*** (0.022)	-0.139*** (0.041)
LAG(ERROR)	Lagged error index (*t* -1)	--	--	--	0.311*** (0.035)	0.431*** (0.062)
N		654	654	654	603	501

[a] Coefficient and standard errors are all long-run effects with the exception of the lagged error. Fixed state effects and state time trends are not shown. Standard errors are in parentheses. See Appendix C for details and calculations.

[b] Model 2a estimates state-specific autocorrelation coefficient. Model 2b estimates a constant autocorrelation coefficient for all States.

For the post-PRWORA period, the estimated effort elasticity ranged from -0.132 to -0.342, with the partial adjustment and Arellano-Bond models showing substantially larger elasticities (in absolute value). Compared with the elasticities for the pre-PRWORA period, the post-PRWORA estimates from the fixed effects and Prais-Winsten models were 51.8 to 56.7 percent smaller (in absolute value), while the estimates from the partial adjustment and Arellano-Bond models were, respectively, 29.8 percent and 9.3 percent smaller (in absolute value). These estimates reflect the combined effects of the EFFORT and PEFFORT variables.[19] Although the post-PRWORA elasticities were less precisely estimated, due largely to the short period (1997-2001) covered by the data, the pre-PRWORA elasticity estimates were statistically different from the post-PRWORA estimates, except for the partial adjustment and Arellano-Bond models. Thus, there is strong and consistent evidence that an increase in effort reduces the error index, and there is also evidence that the magnitude of the effect was probably smaller in the post-PRWORA era (1997-2001)

The model indicates that a combination of changes in the FSP had opposite effects on the elasticity of error with respect to effort. On the one hand, there was a reduction in the absolute value of the slope of the line representing error as a function of effort, thus reducing the absolute value of the elasticity. On the other hand, other factors that reduced the error index (as discussed below) shifted the line down, thus increasing the absolute value of the elasticity (since the same number of units of change in error represented a larger proportional change).[20] The estimates indicate that the effect on the elasticity of error represented by PEFFORT was greater than that of the other effects.

At the national level there was both an increase in effort and a decline in error from 1998 to 2001. During the same period the observed elasticity of error to changes in effort was less in absolute value than before 1996. While we do not have clear evidence of the reasons for these changes, we suggest three alternative explanations below.

One potential explanation is that more effort may have been in fact expended to achieve a given level of accuracy. As noted earlier, we hypothesized that the challenges of implementing PRWORA absorbed staff time that could otherwise have been spent on preventing and detecting errors. It is also plausible that there were lags in the adjustment of staffing to declining FSP caseloads, and that the incremental effort per FSP household was not as focused on error prevention and detection as the previous effort. (This effort was not necessarily wasted, because it might have been focused in improving timeliness or access.) The combination of increased effort without a corresponding reduction in error would help explain the observed decline in the elasticity of error with respect to effort.

An second, alternative explanation is that more of the actual effort may have been charged to the FSP in the post-PRWORA period, so that the observed level of effort for a given level of error was greater. As discussed in an earlier section of this chapter, changes in cost allocation rules resulted in more shared costs for FS/TANF cases being allocated to the FSP, and so States had to spend more FSP dollars per

---

[19] Our estimates of post-PRWORA elasticities were less precise, but all models yielded estimates that were statistically different from zero (*p* <.1 or less).

[20] The estimated elasticity was inversely proportional to the average observed error index for the period for which it was estimated. Thus, all variables that contributed to the reduction in the error index in the post-PRWORA period had the effect of increasing the absolute value of the elasticity.

Table 10

Effort Elasticities in the Pre and Post-PRWORA Periods[a]

	(1) Fixed Effects	(2a) Prais-Winsten (ρ_i)	(2b) Prais-Winsten (ρ)	(3) Partial - Adjustment[b]	(4) Arellano Bond[b]
Pre-PRWORA Elasticity	-0.305 (0.052)	-0.276 (0.060)	-0.300 (0.059)	-0.325 (0.069)	-0.377 (0.116)
Post-PRWORA Elasticity	-0.132 (0.062)	-0.133 (0.069)	-0.142 (0.078)	-0.228 (0.080)	-0.342 (0.147)
N	654	654	654	603	501

[a] Standard errors are found in parentheses.
[b] Long-run effects were estimated for models using lagged error.

household (in real terms) to produce the same output. Thus, the changes in cost allocation reduced the elasticity of error with respect to effort.

Although our analysis did not find clear support for a non-linear model of effort and error, we do not entirely reject this third alternative explanation. It is reasonable to suppose that there is some lower bound to error rates that are realistically attainable, and that decreases in error rates below a certain level require more effort. It is possible that the inclusion of additional years of data would make such a non-linear relationship more apparent.

Factors other than the increase in the measure of effort may have been responsible for the decline in error in the post-PRWORA period. Below, we discuss the relationship of other known covariates to error and the factors that may have contributed to this trend. We also discuss the possibility that the decline in error was partly due to unobservable factors not captured by the known covariates. These factors may have been at least partially incorporated in the fixed state effects or the state time trends.

In the following discussion, we further consider the possible reasons for the change in the relationship of effort to error that is indicated by the model parameters and the elasticity estimates. The interpretations are somewhat speculative because of the limited information and the many factors that may have influenced the trends in error rates. It is important to note that the estimated effect of effort before and after PRWORA is conditional on the other variables in the models, so their effects must be noted when evaluating the estimated effects of effort.

Welfare Reform (PRWORA)

The models show the effects associated with the enactment and implementation of PRWORA in three parts: through the interacted variables PEFFORT and PTANF, and through the PRWORA indicator.

- As previously described, the positive value of PEFFORT means that a given level of certification effort had a reduced effect on the error index in the post-PRWORA period.
- The positive value of PTANF means that States with more FSP households receiving TANF had relatively higher error indexes in the post-PRWORA period.

- The coefficient on the PRWORA indicator is not meaningful in itself because it shows only what the effect of PRWORA would be in a state in which both the fraction of cases on TANF and effort were zero.

To calculate the full effect of PRWORA for any combination of certification effort and TANF participation, one must sum the negative value of the coefficient of the PRWORA indicator and the State-specific effects related to the interaction terms

Percent of FSP Households with AFDC/TANF

Contrary to expectations, the models generally indicated that, before PRWORA, the percent of FSP cases receiving AFDC did not have a significant relationship to the error index. This result was somewhat surprising, because we hypothesized that FSP agencies had an advantage in processing AFDC or TANF cases. The rationale was that the AFDC/TANF benefit was known with certainty, so the potential error in the estimate of total income was less than for households where less readily-determined sources of income made up more of the total. The lack of an effect for the TANF variable was also contrary to our hypothesis that the sharing of costs between the FSP and the AFDC/TANF program might reduce the amount of error with a given level of effort allocated to the FSP.

On the other hand, the interaction of PRWORA with the percentage receiving TANF (the PTANF variable) was significantly and positively associated with the error index, i.e., States with more FSP households receiving TANF had higher error indexes in the post-PRWORA period. The fixed state effects were expected to control for persistent differences among States commonly associated with high AFDC/TANF participation, such as relative levels of urbanization and median income. Thus, the fixed state effects reduced the likelihood that such underlying differences among States confounded the effect of PTANF.

Combined Effect of Variables Related to PRWORA

The net effect of PRWORA was identified as the combined effect of three variables in the models: PEFFORT, PTANF, and the PRWORA indicator. Thus, the net effect of PRWORA depended on a State's level of EFFORT and TANF during the post-PRWORA period—greater for some States than for others. At the mean values of EFFORT and TANF in the post-PRWORA period, the net effect of PRWORA was an increase of 1.6 percentage points in the error index, relative to what it would have been in the absence of PRWORA. The algebraic effect of PRWORA on the error index was smaller for States with a below-average percentage of FSP households receiving TANF or a below-average level of certification-related effort. In fact, for States with the percentage receiving TANF at less than 50 percent of the mean for the post-PRWORA period (i.e. around 12.5 percent) and an average level of certification effort, the net effect of PRWORA on the error index turned negative. Relatively few States reached this level; among the States, the lowest value for the unweighted mean percent with TANF was 20 percent in 2001. For States with the TANF percentage at 50 percent above the mean for the period (about 37.5 percent), the net effect of PRWORA was an increase in the error index of 3.3 percent, more than twice the effect at the mean value.[21] This net effect was, however, smaller than what would be predicted solely on the basis of the effect associated with the PTANF variable.

[21] In the Arellano-Bond model, the PRWORA variable had a similar but non-significant long-run effect. This model is the least efficient due to the loss of information and use of instrumental variables, but it is has the strongest controls against bias in parameter estimates due to effects of lagged error.

It is important to note that, as a group, States experienced a substantial decline in the percent of FSP households receiving TANF. Therefore, the positive effect of PRWORA on the error index through this variable diminished over time. More generally, for each State, the balance of effects of PRWORA on the error index varied from year to year, depending on the values of the TANF and EFFORT variables.

The results may be interpreted as showing that some changes in FSP operations associated with PRWORA implementation had the effect of reducing the level of error, while other changes had the opposite effect. Below, we discuss the potential explanations for these offsetting effects.

Explanations for Effects Related to PRWORA

There are two potential explanations for the post-PRWORA association of a higher percentage of FSP households receiving TANF with a higher error index (i.e., the positive coefficient for PTANF). First, one or more factors in the post-PRWORA environment may have made FSP-TANF cases more prone to FSP errors than other FSP cases. This interpretation is consistent with a hypothesis that implementation of TANF was more disruptive to FSP operations in States with high percentages of FSP cases receiving TANF. A variant of this interpretation is that changes in TANF rules, which often were not matched by changes in FSP rules, had the effect of introducing new possibilities for FSP errors for FS-TANF cases. These two variants are not mutually exclusive.

A second interpretation is that States with the largest decreases in AFDC/TANF caseload also undertook aggressive measures to reduce FSP errors than other States, i.e., that PTANF was negatively correlated with the error rate but proxied for one or more omitted variables. The models controlled for differences in the level of certification-related effort and in the assignment of certification periods. As discussed in more detail below, we did not have data that would explicitly control for other ways in which States may have changed their FSP operations to increase certification accuracy. Thus, we cannot rule out this interpretation.[22]

Turning to the negative component of the PRWORA effect on the error index, there are two potential explanations. The first explanation focuses on the possible effects of PRWORA, while the second explanation involves FSP changes that coincided with PRWORA.

One possible explanation for a negative effect on the error index is that, contrary to some fears at the time, PRWORA implementation had a positive impact on public welfare workers' effectiveness in preventing and detecting errors. During this period, public welfare agencies made a variety of changes in staff training, management, definition of agency mission, and incentives for worker performance. Although many of the changes were driven by PRWORA's goals of increasing clients' rates of employment and reducing their dependence on government assistance, these changes may

[22] It is conceivable that the changes in cost allocation practices after 1996 may have affected the relationship of the percentage of FSP households with TANF to the error index. After PRWORA, States were required to allocate shared costs to all benefiting programs, so the percent of FSP households receiving TANF had less impact on the effective output of a given level of effort allocated to the FSP. This interpretation implies that PTANF, not PEFFORT, captures the effect of the post-PRWORA changes in cost allocation. It is more plausible, however, that the effect of cost allocation changes is captured by PEFFORT, because the magnitude of the change in cost allocation varied across States and was not necessarily related to the percent of FSP cases with TANF. The lack of effect for the TANF variable supports this argument.

have had a beneficial effect on workers' morale and productivity. Declining TANF and FSP caseloads also gave FSP managers an opportunity to increase the emphasis on error reduction.

Another explanation is that the implementation of PRWORA coincided with changes in FNS and State FSP policies and practices that were intended to reduce errors. As noted in Chapter Four, FNS strengthened the financial incentives for States to reduce error rates in a series of steps, starting with settlements regarding outstanding liabilities in 1993. In response to these incentives, States changed their rules regarding reporting and recertification in ways that reduced the likelihood that a QC review would find an error. We have controlled for one widely recognized practice, the use of short certification periods of one to three months (CM13). During the post-PRWORA period, however, a variety of other options were introduced through waivers and rule changes. As described by Rosenbaum (2000), options such as quarterly reporting had the effect of reducing the likelihood of a QC error by narrowing the scope of recipients' responsibilities to report changes. Even if a household's income has changed and the benefit level does not match the current income, there is no error if the household is not required to report the change in the month that is reviewed.[23]

Ideally, the models of error would include variables for quarterly reporting and other practices that reduced the State's exposure to QC errors. No annual data on State adoption of these practices were available, however. Furthermore, some error reduction practices did not require rule changes (e.g., increased monitoring of error rates at the local office level or even the worker level, as described by CBPP, 2001). Thus, we cannot separate the effect of these error reduction practices from other changes occurring after the adoption of PRWORA.[24]

To summarize the preceding discussion, the models capture three distinct effects on error associated with PRWORA.

- During the post-PRWORA period, there was a reduction in the elasticity of error with respect to reported effort.
- The post-PRWORA period also had a pattern in which States with higher percentages of FSP households receiving TANF had higher error rates (all else equal).
- The net effect of PRWORA on the error index was smaller than these effects alone would predict.

The effects associated with PRWORA may have resulted from three types of changes in the FSP and in the operations of public welfare agencies during this period:

---

23 Rules permitting this form of quarterly reporting were issued in July 1999.

24 We explored but eventually rejected models that included policies that PRWORA or previous waivers from AFDC policy permitted States to adopt, such as time limits for AFDC/TANF receipt, earnings disregards, and sanctions policies for violations of AFDC/TANF work requirements. These policies were shown by Kornfeld (2002) to affect FSP participation and, therefore, might affect the composition of the FSP caseload in ways that the available data would not identify. Since these policies were generally intended to increase work or reduce the duration of AFDC/TANF participation, we did not have clear theoretical rationale for how they would affect error rates independently of the percent of FSP households with earnings and the percent with AFDC/TANF, both of which were variables in the model. If such policies did have an additional effect on error rates, their effect may be part of the overall effect of PRWORA that we observed.

- The transition from AFDC to TANF, which entailed changes in both the rules for cash assistance and the environment in which public welfare workers operated
- FNS and State initiatives to reduce FSP errors, through changes in rules and program operations
- The changes in cost allocation rules and practices that resulted in varying increases in the FSP's share of common certification costs for FSP cash assistance households.

The available data were insufficient to determine the relative influence of these three types of changes, each of which had multiple dimensions. As discussed above, we believe that the most convincing explanation for the change in the response of error to reported effort is that it was related to changes in cost allocation rules, though other factors may have contributed. We also believe that the negative component of the effect on error index is at least in part attributable to FSP error-reduction policies other than the shortening of certification periods (represented by the percent of FSP households with 1 to 3-month certification periods, i.e., CM13). We cannot determine from the available data whether the transition from AFDC to TANF had positive or negative effects on FSP errors.

Caseload Characteristics

The preceding results were obtained after controlling for the effects of several important characteristics of FSP households on error rates. These variables were EARNINC, SSINC, and SINGLEPAR. We found the following results for these variables:

- The percent of FSP households reporting earned income (EARNINC) had a positive and highly significant effect on the error index, as expected, with a larger estimated effect from the Arellano-Bond model than the others. Thus, the decline in the error index in the late 1990's was achieved despite the fact that increasing work force participation among FSP recipients exerted upward pressure on the error index.
- The percent of FSP households with Social Security or SSI income (SSINC) had a negative effect on the error index, as expected, and the coefficient was significant in all models except the Arellano-Bond model. This proportion grew during the late 1990's, so this was another factor underlying the decline in the error index.
- The percent of FSP households with children headed by a single adult (SINGLEPAR) had a negative effect on the error index, as expected, but the coefficient was not significant. Thus, it did not appear that the number of adults in the household had an effect on the probability of error, after controlling for receipt of AFDC/TANF, Social Security/SSI income, and earnings.[25]

[25] While it was typical for single-parent FSP households to receive AFDC during the pre-PRWORA period, this association was weaker in the post-PRWORA period, so it is not likely that colinearity is the reason for not finding a significant effect for SINGLEPAR.

Other Independent Variables

Among the other independent variables in the model, only the percentage of FSP households with one to three-month certification periods (CM13) had a significant effect on the error index. The highly significant negative effect of short certifications confirmed the rationale for this practice and the findings of other studies (Kabbani and Wilde, 2003). Neither the percentage of FSP households using EBT (PCTEBT) nor the unemployment rate (FYUN0) had a significant effect on the error index, even after adjusting for the effect of the lagged error index. Kabbani and Wilde (2003) also did not find significant effects for these variables but considered them necessary parts of their model of payment error rates.

Lagged Error Index

The results indicate a highly significant, positive relationship between the lagged error index and the current error index. This finding suggests that, when choosing among the models, the models that incorporate this variable are preferable.

Nevertheless, there is little evidence that the controlling for the effects of lagged error had any significant effect on other parameter estimates, with some exceptions. Comparison of the results across the models indicates a very high degree of consistency in the results. The partial adjustment model yielded very similar results to the models lacking the lagged error variable, with the same variables showing significant effects and parameter estimates generally within one standard error of those of the other models. Most of the parameter estimates from the Arellano-Bond model were also within one standard error of those of all other models, but there were two exceptions. First, the Arellano-Bond model estimated a larger effect of effort on error (in absolute terms), and the difference in estimates between the Arellano-Bond and the Prais-Winsten model was more than one standard error of the former (though less than the sum of the standard errors of the two). This difference is not large enough to affect the overall conclusions or give more credibility to the Arellano-Bond model than any of the others. Second, the Arellano-Bond model yielded an estimate of the effect of percent with earned income (EARNINC) that was two standard errors larger than the estimates from the three models that did not control for lagged error. Since the partial adjustment model also yielded a larger estimated effect for EARNINC than the three models without lagged error, it seems likely that these three models understate the effect of this variable.

Thus, we do not see a convincing reason for preferring the estimates of the Arellano-Bond model to those of the other, more similar models. On balance, given the overall similarity of results, we conclude that the statistical inefficiency of the Arellano-Bond model was most likely the reason that this model did not yield a significant effect for PRWORA or a significant difference in the elasticity of error with respect to effort between the two periods.